骚扰阿蚊线粒体全基因组序列测定和分析

李旭东, 陈　斌*

(重庆师范大学昆虫与分子生物学研究所, 媒介昆虫重庆市重点实验室, 动物生物学重庆市重点实验室, 重庆 401331)

摘要:【目的】测定和分析骚扰阿蚊 *Armigeres subalbatus* 线粒体全基因组序列,并在线粒体基因组水平探讨阿蚊属 *Armigeres* 在库蚊亚科(Culicinae)中的系统发育地位。【方法】经 PCR 扩增和序列测定,首次得到骚扰阿蚊线粒体基因组序列;对其核苷酸组成和结构特点进行分析;基于蛋白质编码基因核苷酸序列,采用最大似然法(ML)和贝叶斯法(BI)构建库蚊亚科 8 个种的系统发育关系。【结果】骚扰阿蚊线粒体基因组全长 14 891 bp(GenBank 登录号: KY978578),包含 37 个基因,其中含 13 个蛋白质编码基因、22 个 tRNA 基因和 2 个 rRNA 基因,各基因位置、排列顺序与蚊科已知物种的一致;基因组碱基组成具有明显的偏好性,全基因组 AT-skew 为正值,GC-skew 为负值;13 个蛋白质编码基因的起始密码子除 *COI* 使用 TCG 外,其余均为 ATN,终止密码子除 *COII* 使用不完全的 T 外,其余均为 TAA;22 个 tRNA 基因中除 $tRNA^{Ser(AGN)}$ 缺失 DHU 臂,其余均可形成典型的三叶草式二级结构。基于库蚊亚科 8 个种的线粒体基因组系统发育关系为库蚊属 *Culex* + (阿蚊属 *Armigeres* + 伊蚊属 *Aedes*)。【结论】分析库蚊亚科的线粒体基因组系统发育关系发现,阿蚊属 *Armigeres* 与伊蚊属 *Aedes* 亲缘关系较其与库蚊属 *Culex* 更近,这与传统的形态分类学结果相吻合。

关键词: 库蚊亚科; 阿蚊属; 骚扰阿蚊; 线粒体基因组; 系统发育

中图分类号: Q969　　**文献标识码:** A　　**文章编号:** 0454-6296(2018)01-0114-08

Sequencing and analysis of the complete mitochondrial genome of *Armigeres subalbatus* (Diptera: Culicidae)

LI Xu-Dong, CHEN Bin* (Chongqing Key Laboratory of Vector Insects, Chongqing Key Laboratory of Animal Biology, Institute of Insect and Molecular Biology, Chongqing Normal University, Chongqing 401331, China)

Abstract:【Aim】To sequence and analyze the complete mitochondrial genome of *Armigeres subalbatus*, and to explore the phylogenetic position of *Armigeres* in Culicinae based on mitochondrial genome sequences.【Methods】The nucleotide sequence of the mitochondrial genome of *Ar. subalbatus* was determined by using PCR amplification and DNA sequencing. The general features and base composition of *Ar. subalbatus* mitochondrial genome were analyzed, and the phylogenetic relationship of *Ar. subalbatus* with seven other species of Culicinae was constructed by using maximum likelihood method (ML) and Bayesian inference (BI) based on the nucleotide sequences of concatenated protein-coding genes (PCGs).【Results】The complete mitogenome of *Ar. subalbatus* is 14 891 bp in length (GenBank accession no.: KY978578), and contains 37 genes, including 13 PCGs, 22 tRNAs and 2 rRNAs. Its gene organization is consistent with those of the known mitogenomes of Culicidae. The mitogenome has a clear bias in nucleotide composition with a positive AT-skew and a negative GC-skew. All 13 PCGs use

基金项目:"两江学者"计划专项经费(20136666); 国家自然科学基金项目(31672363, 31372265); 国际原子能机构 CRP 项目(18268); 国家科技基础性工作专项重点项目(2015FY210300); 广西科技重大专项(桂科 AA17129002)

作者简介: 李旭东, 男, 1987 年 3 月生, 山东烟台人, 硕士, 助理实验师, 研究方向为昆虫分类学, E-mail: 308234091@qq.com

* 通讯作者 Corresponding author, E-mail: bin.chen@cqnu.edu.cn

收稿日期 Received: 2017-07-30; 接受日期 Accepted: 2017-11-21

the ATN as the initiation codon except *COI* with TCG, and TAA as the termination codon except *COII* with the incomplete T. All tRNAs could form the typical cloverleaf structure except tRNA$^{Ser(AGN)}$, whose DHU-arm is lacking. The phylogenetic relationship of eight species within Culicinae based on the mitochondrial genomes presents as *Culex* + (*Armigeres* + *Aedes*). 【Conclusion】 The phylogenetic relationship of Culicinae constructed based on the mitochondrial genomes shows that *Armigeres* is closer related to *Aedes* than to *Culex*, which is consistent with the result of the traditional morphology-based taxonomy.

Key words: Culicinae; *Armigeres*; *Armigeres subalbatus*; mitochondrial genome; phylogeny

根据内共生理论,动物体内的线粒体的祖先为一种好氧性的 α 变形菌,其被真核生物吞噬后,经过漫长的共生过程逐步演化为真核生物体内线粒体(Behura *et al.*, 2011)。作为一种半自主性细胞器,其基因组能够在核基因控制下自主复制,且具有母系遗传、进化速率快、种间多态性大及易于 PCR 扩增等特点(魏书军和陈学新, 2011; Cameron, 2014),已经在分子进化分析、种群遗传结构分析、系统发生学分析等方面得到广泛应用(Brown *et al.*, 1979; Ballard and Whitlock, 2004)。不同物种线粒体基因组大小不一,但对于大部分昆虫而言,其线粒体基因组长度约为 13 ~ 19 kb 左右,含有 37 个基因,包括 13 个蛋白质编码基因(PCGs)、2 个核糖体 RNA 基因(rRNA)、22 个转运 RNA 基因(tRNAs)和一个长的非编码区(又称为控制区或 AT 富含区)(Boore, 1999)。

蚊科昆虫与人类关系密切,能够传播疟疾、寨卡、黄热病和登革热等。据世界卫生组织报道,2015 年全世界有 429 000 人死于疟疾(其中约 303 000 人为 5 岁以下儿童)(World Health Organization, 2016)。控制以上传染病的关键在于控制它们的传播媒介(蚊虫),这依赖于对蚊虫的生物学、遗传学及基因组学的相关基础性研究。阿蚊属 *Armigeres* 蚊虫的系统发育关系尚不清晰,线粒体基因作为分子标记广泛应用于分子鉴定、系统发生学及种群遗传学研究(Wei *et al.*, 2010; Zhao *et al.*, 2010; Kim *et al.*, 2011),因此对于阿蚊属蚊虫线粒体基因组的研究也显得非常重要。

骚扰阿蚊 *Armigeres subalbatus* 在分类学上隶属于蚊科库蚊亚科(Culicinae)阿蚊属(Harbach, 2007)。骚扰阿蚊为阿蚊属中的优势种,主要分布于东洋区与古北区,在我国除东北、新疆及内蒙古等高纬度地区外均有分布(李旭东等, 2013)。成虫吸血凶猛,人和畜血兼吸,且能传播班氏丝虫病,是阿蚊属中最主要的骚扰性吸血蚊种(Tanaka *et al.*, 1979)。本研究首次测得骚扰阿蚊线粒体全基因组序列,并在线粒体基因组水平对其系统发生关系进行探讨,结合已发表的库蚊科 7 种线粒体基因组序列分析了各种属之间的系统发育关系,为后续进一步在线粒体基因组水平分析蚊科各种属间的系统发育关系提供基础。

1　材料与方法

1.1　样本及其 DNA 提取

骚扰阿蚊样本采集于云南省勐腊县的野外畜舍。单头蚊虫样本经乙酸乙酯毒瓶处死后,100% 酒精浸泡密封保存于 −80℃ 条件下。取其头胸部为实验材料,使用 SDS/蛋白酶 K 裂解法提取总 DNA(邹依霖等, 2015)。

PCR 扩增与序列测定:本研究中的 PCR 引物采用已公布的双翅目昆虫线粒体基因组扩增通用引物(张乃心等, 2013)。实验采用 25 μL 反应体系:模板 DNA 2 μL,上、下游引物(10 μmol/L)各 1 μL, dNTP Mix 3 μL, 10 × PCR Buffer(mg^{2+} free) 2.5 μL, $MgCl_2$(25 μmol/L) 1.5 μL, Taq DNA 聚合酶(5 U/μL) 0.2 μL, ddH_2O 13.8 μL。PCR 扩增程序:94℃ 预变性 1 min; 94℃ 变性 40 s, 48 ~ 55℃ 退火 45 s, 68℃ 延伸 1 min, 35 个循环;72℃ 延伸 10 min。扩增产物经 1% 琼脂糖凝胶电泳检测后,委托深圳华大基因双向测序。

1.2　序列拼接及注释分析

测序所得 DNA 片段使用 DNAMAN(http://www.lynnon.com/)进行序列组装拼接。tRNA 基因的查找定位及其二级结构的预测使用在线版的 tRNAscan-SE Search Server v.1.21(http://lowelab.ucsc.edu/tRNAscan-SE/)(Lowe and Eddy, 1997),其中 2 个 tRNA 基因(tRNASer and tRNAArg)无法在线识别,通过与已知蚊种的线粒体基因组进行同源比对识别定位。13 个蛋白质编码基因及 2 个 rRNA

基因利用 Clustal X(Thompson *et al.*, 1997)与已知蚊种的线粒体基因进行同源比对定位。利用 Mega 5.0 分别计算骚扰阿蚊线粒体基因组中各编码基因的碱基组成、密码子使用频率、AT-skew 和 GC-skew(Tamura *et al.*, 2011)。

1.3 系统发育分析

利用新测得的骚扰阿蚊线粒体基因组的蛋白质编码基因氨基酸,选取 GenBank 中已报道的 7 种库蚊亚科蚊虫的线粒体基因组中 13 个蛋白质编码基因的核苷酸序列,以冈比亚按蚊 *Anopheles gambiae* 为外群,采用最大似然法(maximum likelihood, ML)和采用贝叶斯法(Bayesian inference, BI)分别构建系统发育树。利用 Clustal X(Larkin *et al.*, 2007)对核苷酸序列进行比对后,用 Modeltest 3.7 软件计算得出建树的最适核苷酸替代模型为 GTR + I + G。以 MEGA 5.0 重复计算 1 000 次,构建 ML 系统发育树及算得检验各分支节点的 bootstrap 置信值。运行 MrBayes 3.1.2 软件计算 1×10^6 代,每运行计算 100 代取样保存,计算得到的前 2.5 千代进化树舍弃,以构建 BI 系统发育树。

2 结果

2.1 骚扰阿蚊线粒体基因组结构

骚扰阿蚊线粒体基因组全长 14 891 bp(不含 AT 区)(GenBank 登录号: KY978578)。包含有 37 个基因(13 个蛋白质编码基因、22 个 tRNA 基因和 2 个 rRNA 基因)。基因序列与已报道的其他蚊种相似度高(Beard *et al.*, 1993; Hua *et al.*, 2016)。22 个基因(9 个蛋白质编码基因和13个tRNA基因)位于 J 链上,其余 15 个基因(4 个蛋白质编码基因、9 个 tRNA 基因和 2 个 rRNA 基因)位于 N 链上(图 1)。

骚扰阿蚊线粒体基因组中的基因排列紧密,但也存在基因重叠及间隔现象。其中,存在 12 处长度 1 ~23 bp 的碱基间隔现象和 13 处长度 1 ~7 bp 的碱基重叠现象(表 1)。

2.2 骚扰阿蚊线粒体基因组碱基组成

骚扰阿蚊线粒体基因组碱基组成见表 2。全基因组的 A + T 含量达到 78.8%,明显高于 G + C 含量(21.2%),呈现出明显的 AT 偏向性。其中蛋白质编码基因、tRNA 基因和 rRNA 基因 A + T 含量分别为 77.7%, 80.8% 和 82.8%。骚扰阿蚊整个线粒体基因组不同位置对碱基偏好性具有一定的差异。总的来看,全基因组的 AT-skew 为正值(0.01),

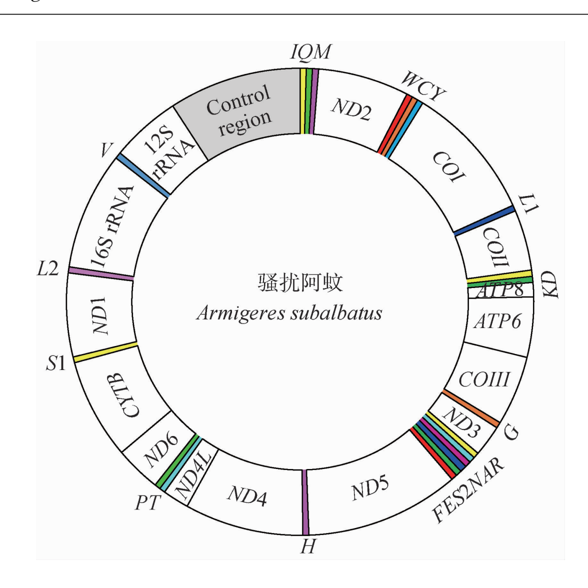


图 1 骚扰阿蚊线粒体基因组结构图

Fig. 1 Structure of *Armigeres subalbatus* mitochondrial genome

图中彩色填充框表示 tRNA 编码基因,未填充的白色框表示蛋白质编码基因和 rRNA 基因,灰色为未能测序的控制区。用黑色和红色字体标记的蛋白编码基因、rRNA 基因及控制区分别位于 J 链和 N 链上。The color-filled blocks indicate tRNAs, while the unfilled white blocks denote protein-coding genes and rRNA gene and grey blocks indicate unsequenced control regions. The protein-coding genes, rRNA gene and control regions with black name are located on majority strand, whereas those with red names are on minority strand.

GC-skew 为负值(-0.20),表明整个基因组更偏好于使用 A 碱基和 C 碱基。从不同的基因来看,tRNA 基因和 rRNA 的 AT-skew 为正值,表明其偏好 A 碱基,其余基因则偏向于使用 T 碱基;蛋白质编码基因、J 链蛋白质编码基因及其第 2 和 3 位密码子、N 链蛋白质编码基因的第 3 位密码子以及 rRNA 基因的 GC-skew 为负值,表明其偏好使用 C 碱基,其余基因 GC-skew 为正值,更偏好使用 G 碱基。

2.3 骚扰阿蚊线粒体基因组蛋白质编码基因

骚扰阿蚊线粒体的蛋白质编码基因中,除了 *COI* 使用 TCG 作为起始密码子外,其余的 12 个蛋白质编码基因均以通用密码子 ATN 作为起始密码子。骚扰阿蚊 7 个基因(*COII*, *ATP6*, *COIII*, *ND5*, *ND4*, *ND4L* 和 *CYTB*)以 ATG 作为起始密码子;*ND1* 以 ATA 作为起始密码,*ND2* 和 *ATP8* 2 个基因以 ATC 作为起始密码子,*ND3* 和 *ND6* 2 个基因以 ATT 作为起始密码子。而在终止密码子的使用上,除 *COII* 基因使用不完全的 T 作为终止密码子外,其余 12 个蛋白质编码基因均以 TAA 作为终止密码子(表 1)。

在大部分动物线粒体基因组中,核苷酸组成的差异性会影响蛋白质编码基因中的氨基酸组成(Foster *et al.*, 1997)。骚扰阿蚊蛋白质编码基因中

也呈现出明显的氨基酸偏好性,20 个氨基酸含量差异明显,其中亮氨酸含量最高(13.2%),谷氨酸含量最低(1.24%)(图 2)。相对同义密码子使用频率(RSCU)是评估同义密码子使用频率的参考值。密码子 UUA/CGA/GGA/UCU 使用频率较高,其中 UUA 最高;CUG/CCG/ACG/CAG/AAC/UGC/AGC/AGG/GUC/GUG/GAG/GGC 使用频率较低,AGG 和 GGC 最低(图 3)。各密码子 RSCU 值相差较大,表明在骚扰阿蚊线粒体基因组中密码子使用频率具有明显的偏向性。

表 1　骚扰阿蚊线粒体基因组组成

Table 1　Composition of *Armigeres subalbatus* mitochondrial genome

基因 Genes	位置(bp) Position	长度(bp) Length	编码链 Coding strand	反密码子 Anticodon	起始密码子 Start codon	终止密码子 Stop codon	基因间的间隔(bp) Intergenic length
$tRNA^{Ile}$	1 – 69	69	J	GAT			0
$tRNA^{Gln}$	67 – 135	69	N	TTG			3(–)
$tRNA^{Met}$	135 – 203	69	J	CAT			1(–)
*ND*2	204 – 1 232	1 029	J		ATC	TAA	0
$tRNA^{Trp}$	1 234 – 1 302	69	J	TCA			1(+)
$tRNA^{Cys}$	1 302 – 1 368	67	N	GCA			1(–)
$tRNA^{Tyr}$	1 369 – 1 434	66	N	GTA			0
COI	1 433 – 2 974	1 542	J		TCG	TAA	2(–)
$tRNA^{Leu(UUR)}$	2 970 – 3 036	67	J	TAA			5(–)
COII	3 039 – 3 723	685	J		ATG	T	2(+)
$tRNA^{Lys}$	3 724 – 3 794	71	J	CTT			0
$tRNA^{Asp}$	3 806 – 3 873	68	J	GTC			11(+)
*ATP*8	3 874 – 4 035	162	J		ATC	TAA	0
*ATP*6	4 029 – 4 709	681	J		ATG	TAA	7(–)
COIII	4 724 – 5 513	789	J		ATG	TAA	15(+)
$tRNA^{Gly}$	5 513 – 5 579	67	J	TCC			1(–)
*ND*3	5 580 – 5 933	354	J		ATT	TAA	0
$tRNA^{Arg}$	5 933 – 5 996	64	J	TCG			1(–)
$tRNA^{Ala}$	5 997 – 6 064	68	J	TGC			0
$tRNA^{Asn}$	6 065 – 6 131	67	J	GTT			0
$tRNA^{Ser(AGN)}$	6 131 – 6 207	77	N	GCT			1(–)
$tRNA^{Glu}$	6 208 – 6 274	67	J	TTC			0
$tRNA^{Phe}$	6 273 – 6 339	67	N	GAA			2(–)
*ND*5	6 347 – 8 090	1 743	N		ATG	TAA	8(+)
$tRNA^{His}$	8 084 – 8 151	68	N	GTG			7(–)
*ND*4	8 155 – 9 498	1 344	N		ATG	TAA	3(+)
*ND*4*L*	9 492 – 9 788	297	N		ATG	TAA	7(–)
$tRNA^{Thr}$	9 791 – 9 855	65	J	TGT			2(+)
$tRNA^{Pro}$	9 856 – 9 921	66	N	TGG			0
*ND*6	9 924 – 10 446	522	J		ATT	TAA	2(+)
CYTB	10 452 – 11 588	1 137	J		ATG	TAA	5(+)
$tRNA^{Ser(UCN)}$	11 587 – 11 652	66	J	TGA			2(–)
*ND*1	11 676 – 12 620	945	N		ATA	TAA	23(+)
$tRNA^{Leu(CUN)}$	12 627 – 12 694	68	N	TAG			6(+)
16S rRNA	12 695 – 14 028	1 334	N				0
$tRNA^{Val}$	14 030 – 14 101	72	N	TAC			1(+)
12S rRNA	14 102 – 14 891	790	N				

“基因间的间隔”一列中,正号表示基因间隔碱基,负号表示基因重叠碱基。In the column of intergenic length, the positive sign indicates the interval base pairs between genes, while the negative sign indicates the overlapping base pairs between genes.

表 2 骚扰阿蚊线粒体基因组核苷酸组成

Table 2 Nucleotide composition of *Armigeres subalbatus* mitochondrial genome

基因序列 Genome sequences	A%	T%	G%	C%	A + T%	G + C%	AT-skew	GC-skew
全基因组 Whole genome	39.7	39.1	8.4	12.7	78.8	21.2	0.01	-0.20
蛋白质编码基因 Protein-coding genes	32.9	44.8	10.8	11.5	77.7	22.3	-0.15	-0.03
J 链上的蛋白质编码基因 Protein-coding genes on J-strand	33.6	43.0	10.1	13.3	77.6	22.4	-0.12	-0.14
第 1 位密码子 1st codon position	34.5	35.5	16.1	13.9	70.0	30.0	-0.01	0.07
第 2 位密码子 2nd codon position	44.7	48.3	1.0	6.0	93.0	7.0	-0.04	-0.71
第 3 位密码子 3rd codon position	21.6	45.2	12.4	20.8	66.8	33.2	-0.35	-0.25
N 链上的蛋白质编码基因 Protein-coding genes on N-strand	32.0	47.7	13.4	6.9	79.7	20.3	-0.20	0.32
第 1 位密码子 1st codon position	31.7	42.9	19	6.4	74.6	25.4	-0.15	0.50
第 2 位密码子 2nd codon position	20.3	50.1	16	13.6	70.4	29.6	-0.42	0.08
第 3 位密码子 3rd codon position	43.9	50	0.8	5.3	93.9	6.1	-0.07	-0.74
tRNA 基因 tRNA	40.5	40.3	11.0	8.1	80.8	19.2	0.01	0.15
J 链上的 tRNA 基因 tRNA genes on J-strand	41.0	39.3	10.0	9.6	80.4	19.6	0.02	0.02
N 链上的 tRNA 基因 tRNA genes on N-strand	39.8	41.6	12.4	6.1	81.5	18.5	-0.02	0.34
rRNA 基因 rRNA	42.2	40.6	6.1	11.2	82.8	17.3	0.02	-0.20

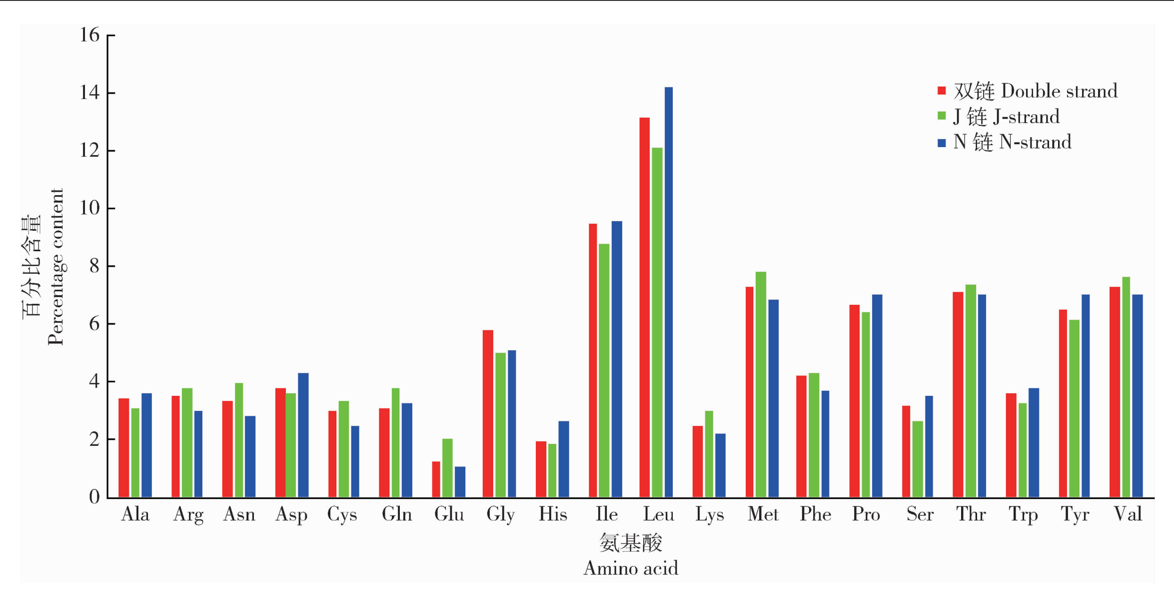


图 2 骚扰阿蚊线粒体基因组上蛋白质编码基因的氨基酸百分比含量

Fig. 2 Percentage content of each amino acid of protein-coding genes in *Armigeres subalbatus* mitochondrial genome

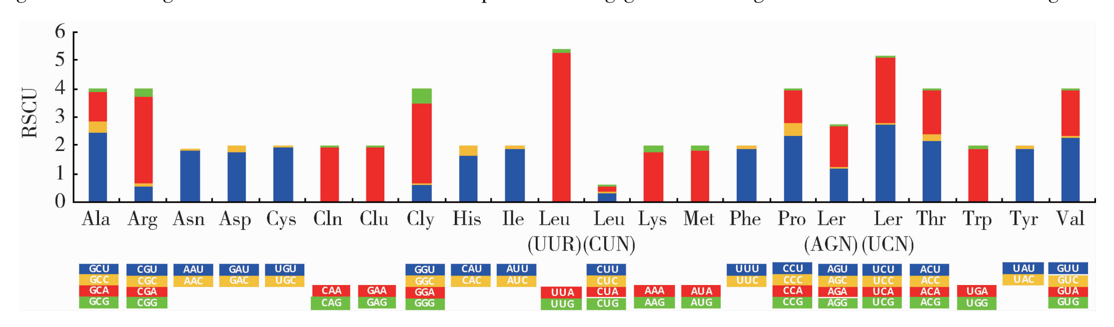


图 3 骚扰阿蚊线粒体基因组相对同义密码子使用频率(RSCU)

Fig. 3 Relative synonymous codon usage (RSCU) in *Armigeres subalbatus* mitochondrial genome

2.4 骚扰阿蚊线粒体基因组 rRNA 基因及 tRNA 基因

骚扰阿蚊线粒体基因组含有 2 个 rRNA 基因，即 srRNA 和 lrRNA，均在 N 链上。lrRNA 位于 $tRNA^{Leu(CUN)}$ 和 $tRNA^{Val}$ 间，长度 1 334 bp；srRNA 在 $tRNA^{Val}$ 和控制区间，长度 790 bp。AT 百分含量分

别是 83.7% 和 81.3% ,低于总体水平。

骚扰阿蚊的 22 个 tRNA 基因大小在 64 ~ 77 bp 之间,总长度为 1 497 bp,在 N 链和 J 链上均有分布。与所有的后生动物一样,骚扰阿蚊的 22 个 tRNA 编码基因中只有 tRNA$^{Ser(AGN)}$ 因缺失 DHU 臂不能形成完整的三叶草式的二级结构(图 4)(Negrisolo, 2011)。

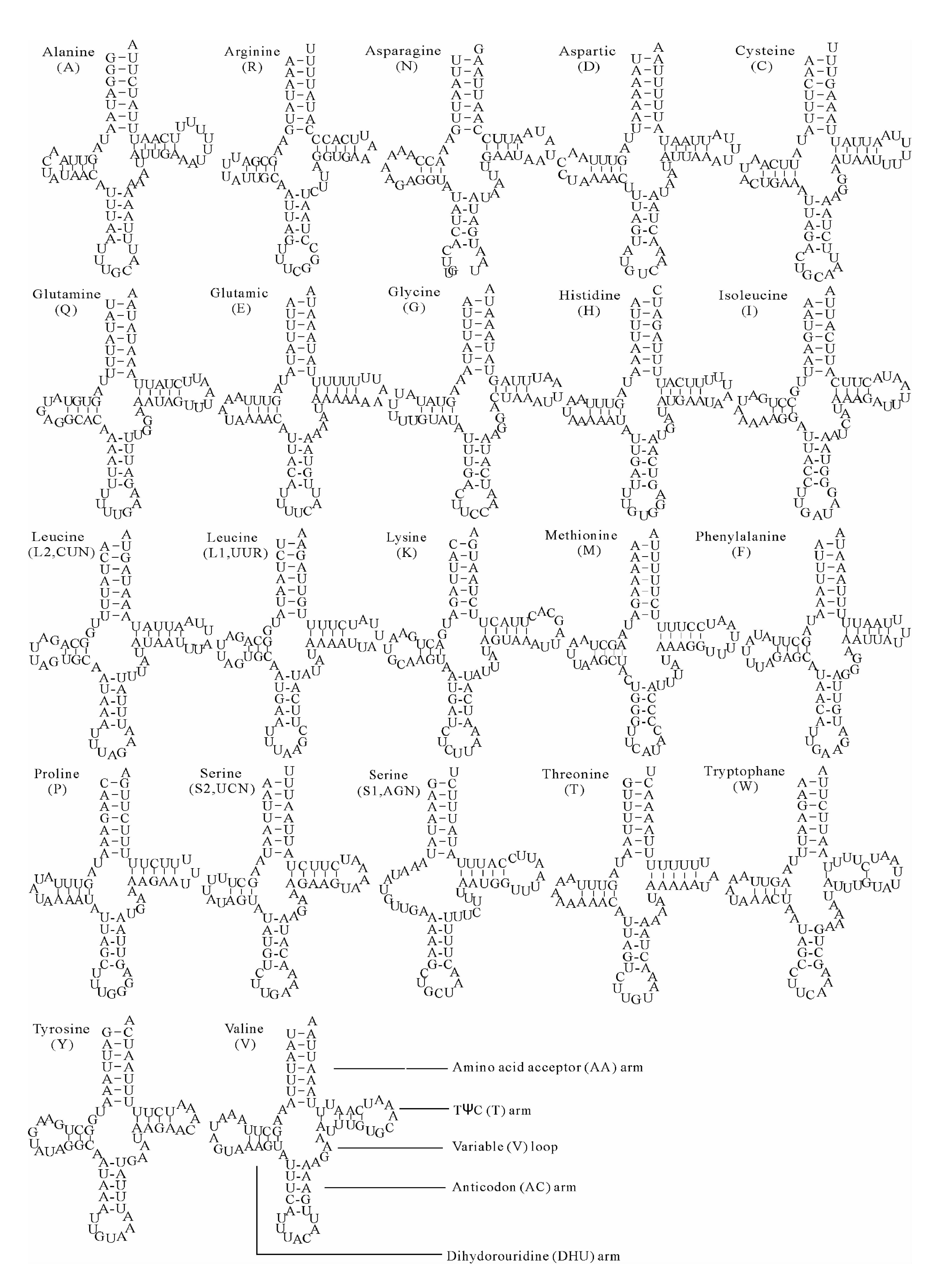


图 4 骚扰阿蚊线粒体基因组 tRNA 基因二级结构

Fig. 4 Predicted secondary structures of tRNA in *Armigeres subalbatus* mitochondrial genome

2.5 库蚊亚科的系统发育关系

以按蚊亚科(Anophelinae)中的冈比亚按蚊 *Anopheles gambiae* 为外群,使用已报道的7种库蚊亚科蚊虫及本研究测得的骚扰阿蚊的线粒体基因组的13个蛋白质编码基因核苷酸序列,分别采用最大似然法(ML)和贝叶斯法(BI)构建库蚊亚科昆虫的系统发育树。结果显示两种建树方法得到的拓扑结构一致,均将库蚊亚科分为2支,一支包含所有的库蚊属 *Culex* (boostrap 值 100%),另一支包括了伊蚊属 *Aedes* 和阿蚊属 *Armigeres* (boostrap 值 90%)。在两个系统发育树中,骚扰阿蚊均聚集在伊蚊属 *Aedes* 分支系,与伊蚊属 *Aedes* 组成姐妹关系,且置信值均大于80%。在本研究中,骚扰阿蚊与库蚊属 *Culex* 蚊虫的遗传距离为0.08 (0.03 +0.05),远大于骚扰阿蚊与伊蚊属 *Aedes* 的遗传距离(0.00),表明骚扰阿蚊和伊蚊属 *Aedes* 蚊虫有更近的亲缘关系,特别是与伊蚊属 *Aedes* 中的纷蚊亚属 *Finlaya* 关系最近。3个属之间的系统发生关系为库蚊属 *Culex* +(阿蚊属 *Armigeres* +伊蚊属 *Aedes*),各属内亚属均聚为一支,且置信值较高,系统发育关系可信度较高。这与现沿用的形态分类学结果相一致,很好地支持了 Harbach 对蚊科的最新分类系统(Harbach, 2007)。

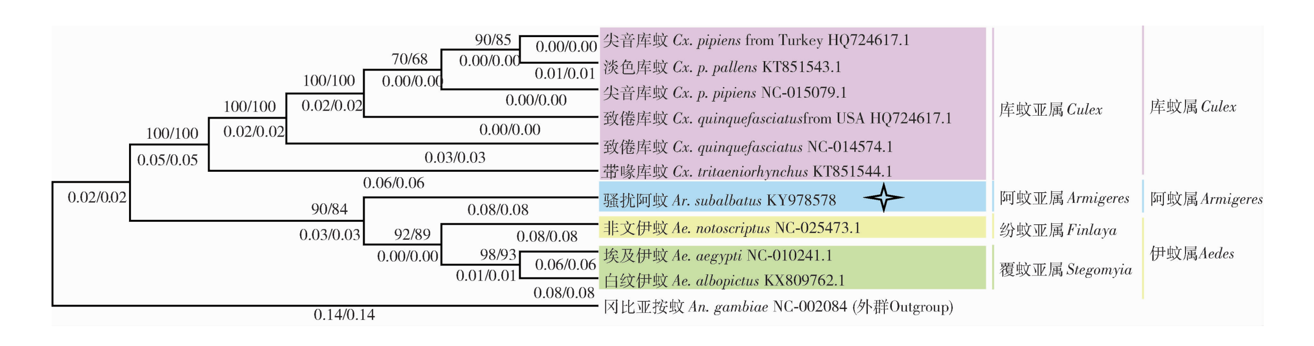


图5 基于线粒体基因组的13个蛋白质编码基因核苷酸序列的库蚊亚科系统发育树(最大似然法和贝叶斯法)

Fig. 5 Phylogenetic tree of Culicinae based on the nucleotides of 13 protein-coding genes of mitochondrial genomes (maximum likelihood method and Bayesian interference)

进化枝上方为置信值,下方为遗传距离;红色和蓝色数值分别代表最大似然法和贝叶斯法所计算的结果。The values on top and bottom of the branches represent the confidence values and genetic distance, respectively. Red and blue values stand for the results of the maximum likelihood estimation and Bayesian algorithm, respectively.

3 讨论

线粒体基因作为重要的母性遗传标记被广泛应用于分子进化分析、种群遗传结构分析、系统发生学分析等方面。本研究首次测得阿蚊属蚊虫的线粒体全基因组序列,并对其线粒体基因组进行了注释和分析,丰富了蚊科线粒体基因组的数据信息,为研究蚊科的分子系统发育进化关系提供了数据基础,特别是为研究阿蚊属在库蚊亚科中的系统发育位置奠定了基础;其次,骚扰阿蚊线粒体基因组的结构特点和碱基组成分析对库蚊亚科线粒体基因组水平的比较分析提供了基础数据;最后,基于骚扰阿蚊和9条已知蚊虫的线粒体基因组中13个蛋白质编码基因序列构建了库蚊亚科的系统发育关系,为阿蚊属在库蚊亚科中的分类地位提供了分子依据,也为后续构建蚊科系统发育关系提供了分子数据。

到目前为止,蚊科共包含2亚科,112属,3 557种(Harbach, 2017),已公开发表50种(亚种)蚊虫线粒体基因组数据,线粒体基因组数据仍远远不足,后续需要进一步对蚊科其他种属的线粒体基因组测序分析,丰富蚊科线粒体基因组数据,从而构建蚊科各亚科、属及亚属间的系统发生关系。

参考文献(References)

Ballard JW, Whitlock MC, 2004. The incomplete natural history of mitochondria. *Mol. Ecol.*, 13(4): 729 – 744.

Beard CB, Hamm DM, Collins FH, 1993. The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. *Insect Mol. Biol.*, 2(2): 103 – 124.

Behura SK, Lobo NF, Haas B, Debruyn B, Lovin, DD, Shumway MF, Puiu D, Severson, JR, Nene V, Severson DW, 2011. Complete sequences of mitochondrial genomes of *Aedes aegypti* and *Culex quinquefasciatus* and comparative analysis of mitochondrial DNA fragments inserted in the nuclear genomes. *Insect Biochem. Molec. Biol.*, 41(10): 770 – 777.

Boore JL, 1999. Animal mitochondrial genomes. *Nucleic Acids Res.*, 27

(8)：1767－1780.

Brown WM, George M, Wilson AC, 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA*, 76(4)：1967－1971.

Cameron SL, 2014. Insect mitochondrial genomics：implications for evolution and phylogeny. *Ann. Rev. Entomol.*, 59(1)：95－117.

Foster PG, Jermiin LS, Hickey DA, 1997. Nucleotide composition bias affects amino acid content in proteins coded by animal mitochondria. *J. Mol. Evol.*, 44(3)：282－288.

Harbach RE, 2007. The Culicidae (Diptera)：a review of taxonomy, classification and phylogeny. *Zootaxa*, 1668：591－638.

Harbach RE, 2017. Mosquito taxonomic inventory [EB/OL]. Available at：http：//mosquito taxonomic-inventory. info.

Hua YQ, Ding YR, Yan ZT, Si FL, Luo QC, Chen B, 2016. The complete mitochondrial genome of *Anopheles minimus* (Diptera：Culicidae) and the phylogenetics of known *Anopheles* mitogenomes. *Insect Sci.*, 23(3)：353－365.

Kim MJ, Kang AR, Jeong HC, Kim KG, Kim I, 2011. Reconstructing intraordinal relationships in Lepidoptera using mitochondrial genome data with the description of two newly sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera：Lycaenidae). *Mol. Phylogenet. Evol.*, 61(2)：436－445.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21)：2947－2948.

Li XD, Yan ZT, Fu WB, Chen B, 2013. Taxonomic status and geographical distribution of the genus *Armigeres* (Diptera：Culicidae). *Chin. J. Vector Biol. Control*, 14(5)：377－381. [李旭东，闫振天，付文博，陈斌，2013. 阿蚊属分类研究进展及地理区系分析(双翅目：蚊科). 中国媒介生物学及控制杂志，24(5)：377－381]

Lowe TM, Eddy SR, 1997. tRNAscan-SE：a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, 25(5)：955－964.

Negrisolo E, Babbucci M, Patarnello T, 2011. The mitogenome of the ascalaphid owlfly *Libelloides macaronius* and comparative evolutionary mitochondriomics of neuropterid insects. *BMC Genomics*, 12(1)：221.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, 2011. MEGA5：molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, 28(10)：2731－2739.

Tanaka K, Mizusawa K, Saugstad ES, 1979. A revision of the adult and larval mosquitoes of Japan (including the Ryukyu Archipelago and the Ogasawara Islands) and Korea (Diptera：Culicidae). *Contrib. Amer. Ent. Inst.*, 16：1－987.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, 1997. The CLUSTAL X Windows interface：flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, 25(24)：4876－4882.

Wei SJ, Chen XX, 2011. Progress in research on the comparative mitogenomics of insects. *Chin. J. Appl. Entomol.*, 48(6)：1573－1585. [魏书军，陈学新，2011. 昆虫比较线粒体基因组学研究进展. 应用昆虫学报，48(6)：1573－1585]

Wei SJ, Min S, Sharkey MJ, Achterberg CV, Chen XX, 2010. Comparative mitogenomics of Braconidae (Insecta：Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to holometabolous insects. *BMC Genomics*, 11(24)：371.

World Health Organization, 2016. World Malaria Report 2016. World Health Organization, Geneva.

Zhang NX, Zhang YJ, Yu G, Chen B, 2013. Structure characteristics of the mitochondrial genomes of Diptera and design and application of universal primers for their sequencing. *Acta Entomol. Sin.*, 56(4)：398－407. [张乃心，张玉娟，余果，陈斌，2013. 双翅目昆虫线粒体基因组结构特点及其测序通用引物的设计和应用. 昆虫学报，56(4)：398－407]

Zhao L, Zheng ZM, Huang Y, Sun HM, 2010. A comparative analysis of mitochondrial genomes in Orthoptera (Arthropoda：Insecta) and genome descriptions of three grasshopper species. *Zool. Sci.*, 27(8)：662－672.

Zou YL, Ding YR, Luo QC, Chen B, 2015. The extraction method of mosquito mitogenome. *Chin. J. Vector Biol. Control*, 26(4)：333－336. [邹依霖，丁奕然，罗钱春，陈斌，2015. 蚊虫线粒体基因组 DNA 提取方法. 中国媒介生物学及控制杂志，26(4)：333－336]

(责任编辑：马丽萍)